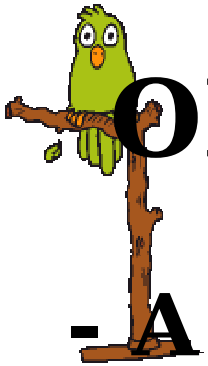
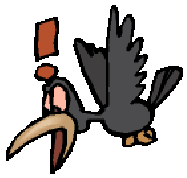




OBJECTIVE VALIDATION

- A Research Initiative Proposal -

S.Y. Harmon
harmon@zetetix.com

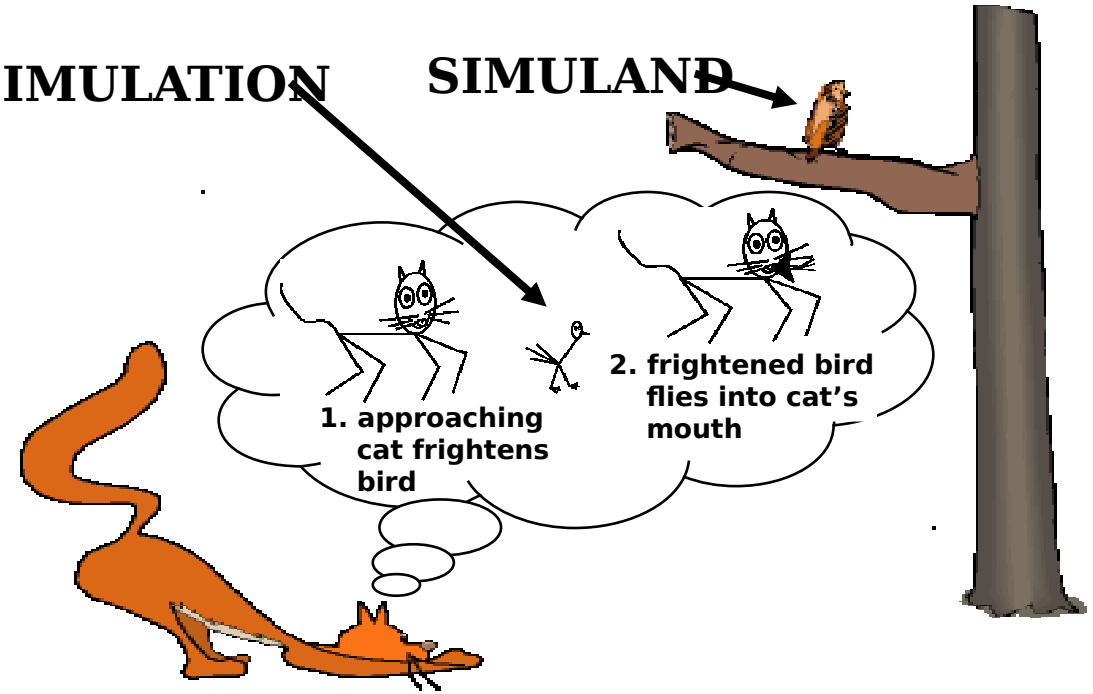
August 2001

(Any other approach to validation is for the birds.)

Zetetix

ALL SIMULATIONS ABSTRACT SIMULAND BEHAVIOR

Everyone knows that, unique among systems, simulations abstractly represent the behavior of something else for some purpose. But, since simulations necessarily omit some of the details about the things they model from



How closely must a simulation resemble its simuland to achieve a particular purpose?"

The ability to answer this question will also address such related questions as "Can a particular existing simulation achieve a purpose for which it was not originally designed?" and "Can a federation of simulations achieve some given purpose?"

Zetetix

VALIDATION ASSESSES SIMULATION FITNESS FOR A PURPOSE

Validation answers this question by assessing the fitness of a simulation for a particular purpose.

The reliability of this assessment and the answers it supplies depend upon the quality of three types of information:

- **Acceptability criteria,**
- **Referent, and**
- **Simulation capabilities.**

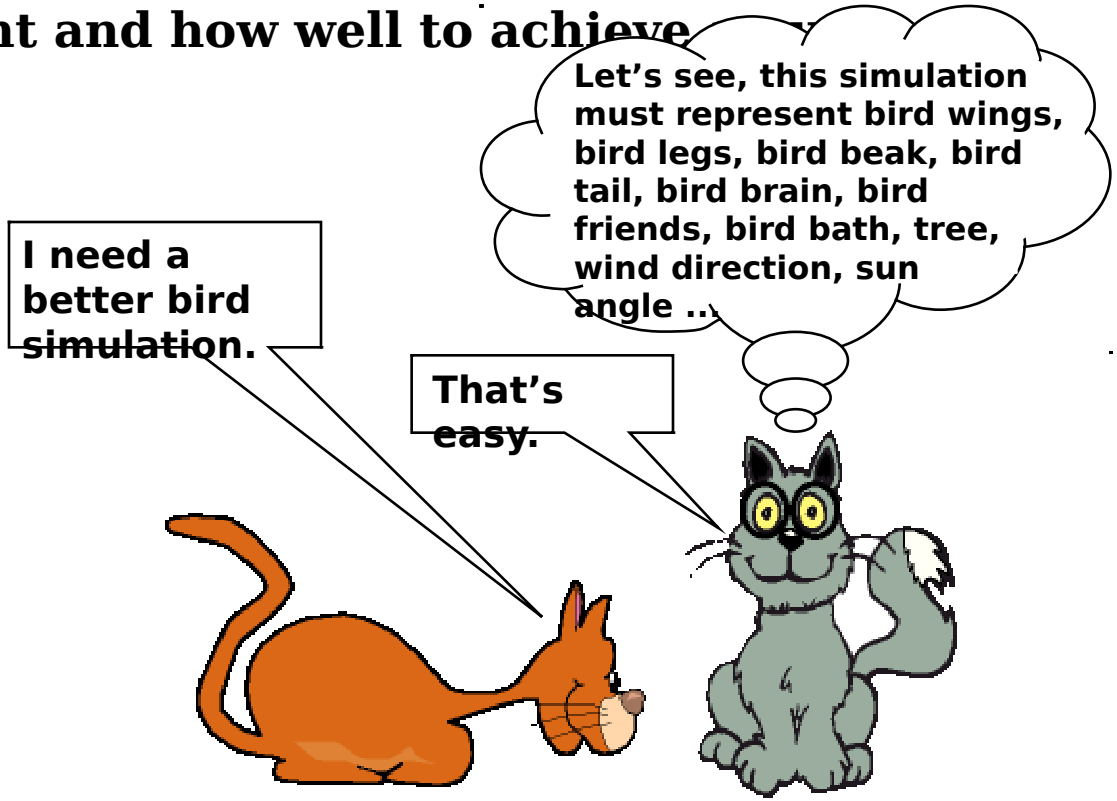
But, getting this information with the quality needed can be challenging, especially for simulations of complex phenomena.

Zetetix

ACCEPTABILITY CRITERIA ELABORATE REQUIRED SIMULATION CAPABILITIES

Acceptability criteria, derived from the user's purpose for a simulation, detail the minimum simulation capabilities needed to achieve that purpose. These criteria define what a simulation must represent and how well to achieve it.

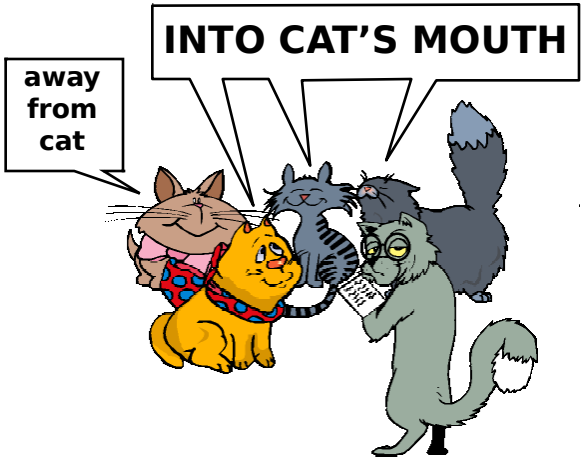
Often subject matter experts (SMEs), working for the user or the developer, derive detailed acceptability criteria for a simulation from very limited user input. These derivations depend upon the SMEs' own subjective opinions of what simulation capabilities the users really need. This subjectivity can decrease the reliability of the resulting criteria and any simulations built to satisfy them.


Let's see, this simulation must represent bird wings, bird legs, bird beak, bird tail, bird brain, bird friends, bird bath, tree, wind direction, sun angle ...
I need a better bird simulation.
That's easy.

Zetetix

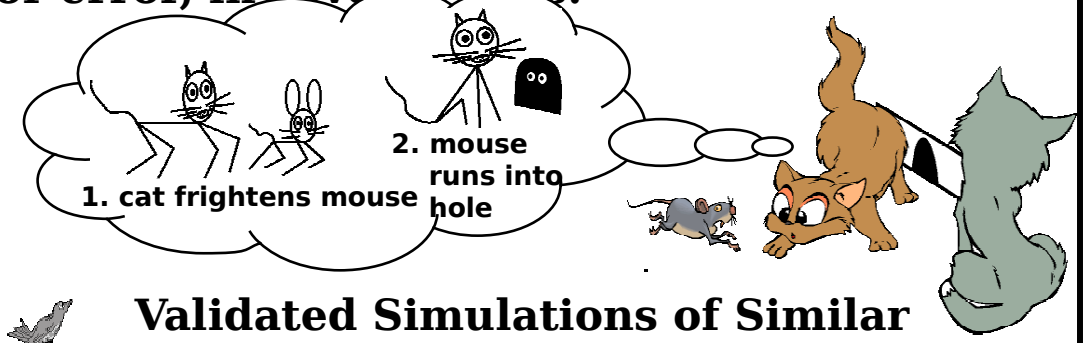
REFERENTS DEFINE THE STANDARDS TO GAUGE ACCURACY

One can construct a referent for phenomena of interest, the best knowledge we have about those phenomena against which to define and measure simulation accuracy or error, in several ways:



SME Opinions

Direct Observations



Validated Simulations of Similar Situations

Theory of Bird Response

Outcome	Probability
No response	0.05
Fly away from cat	0.95
Fly into cat's mouth	0.00

Theory Derived from Observations

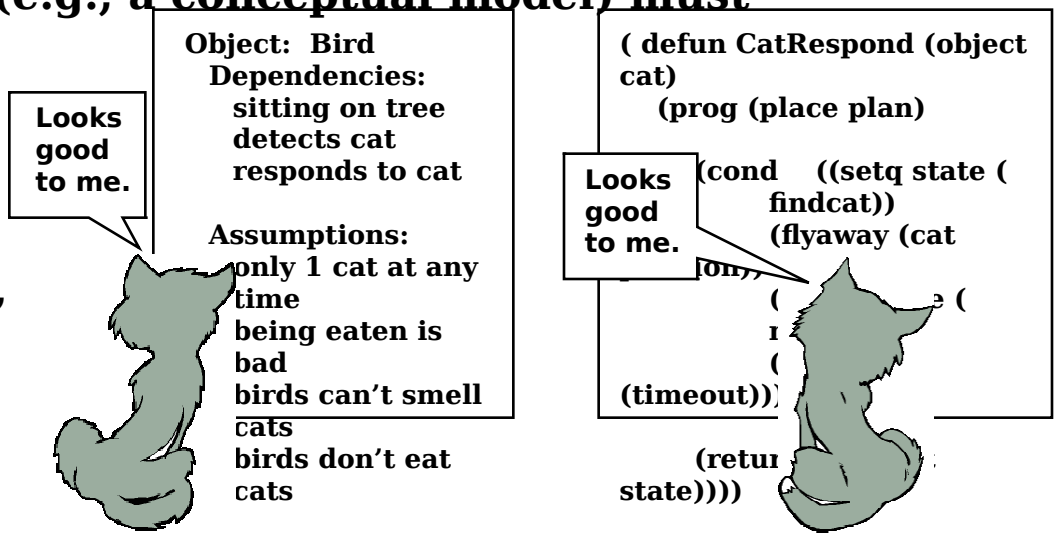
The most commonly used but, by far, the weakest referent knowledge comes from SME opinions.

Zetetix

A SIMULATION'S DESIGN PROVIDES INITIAL VISIBILITY INTO ITS CAPABILITIES

But, for simulation design validation to be reliable and useful for validation, its documentation (e.g., a conceptual model) must

- **Contain sufficiently rich descriptions of representational capabilities,**
- **Describe capabilities in a form accessible to evaluators, and**
- **Elucidate the assumptions underlying all representation design decisions.**



Good Design Documentation **Poor Design Documenta**

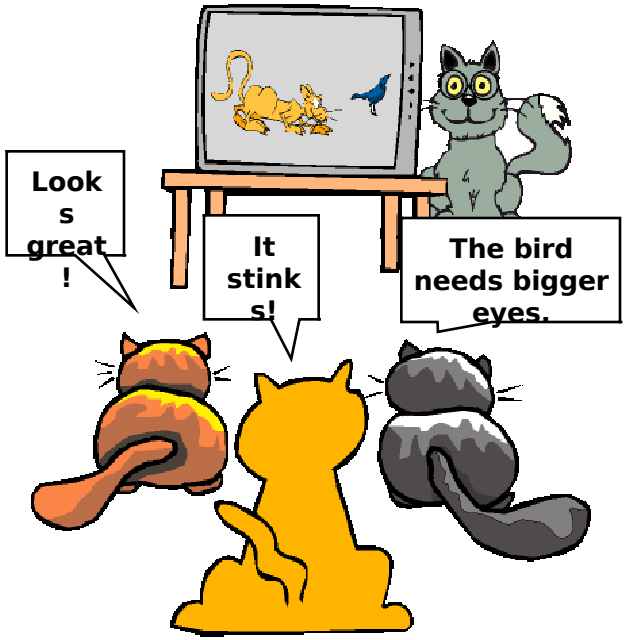
Even scrupulously conducted design reviews incorporate many subjective and poorly characterized elements that weaken design validation assessments.

Validated designs create the only reference point upon which all subsequent verification efforts depend in order for those efforts to contribute substantially to simulation validation and, thus, credibility.

Zetetix

SIMULATION RESULTS SUPPLY THE MOST DIRECT CAPABILITIES VISIBILITY

Actual results produced by a simulation create the last, most important and, often, most accessible depiction of that simulation's capabilities. But,



- **Practical results validation usually cannot exhaustively explore the behavior spaces generated by complex simulations.**
- **Sampling complex behavior spaces for results validation has poorly understood consequences upon assessment reliability and utility.**
- **SME assessments tend to inextricably couple requirements, referent and capabilities information thus severely complicating their interpretation.**
- **SME validity opinions can vary widely, even conflict, and frequently do.**

Even for simulations of well understood physical phenomena, much of existing results validation relies primarily upon the subjective opinions of SMEs.

Zetetix

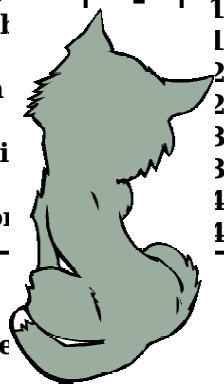
VALIDATION INFORMATION MUST BE EXPRESSED IN COMPARABLE TERMS

Meaningful and, thus credible, comparison of referent, acceptability criteria and capabilities demands their expression in directly comparable terms (e.g., same level of abstraction, resolution, units, coordinate systems).

Acceptability Criteria

1. Bird must fly
2. Bird must walk on limbs
3. Bird must sense cat at a reasonal distance
4. Bird must sit on trees & ground
5. Bird must use bi bath
6. Bird must respor to cat as a real bird would.
7. Bird must show effects of fatigue

?



Simulation Results		
time	cat position	bird position
0	40, 20, 0	0, 0, 5
5	38, 18, 0	0, 0, 5
10	37, 17, 0	0, 0, 5
15	35, 16, 0	0.5, 0, 4.8
20	33, 15, 0	0.5, 0, 4.8
25	33, 13, 0	0.5, 0, 4.8
30	31, 13, 0	0.5, 0, 4.8
35	29, 12, 0	0, 0, 5
40	28, 12, 0	0, 0, 5
45	26, 12, 0	0, 0, 5

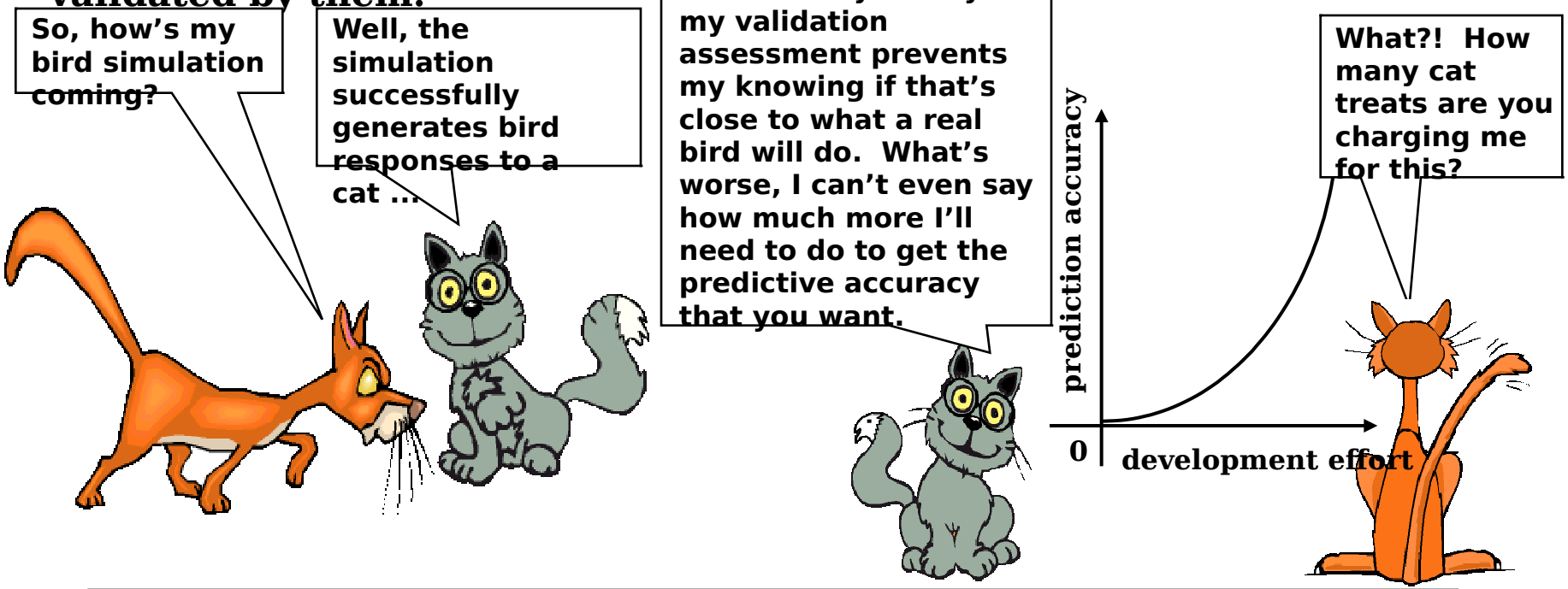
Describing validation information in different terms and at different levels of abstraction, as often happens, necessitates using SMEs to interpret and compare it and, so, introduces considerable subjectivity into the validation assessment thus weakening that assessment even when the contributing information has been

Creating the means to concisely and consistently describe simulation requirements, referents and capabilities calls for a fundamental understanding of the nature of simulation itself, an understanding that currently does not exist.

Zetetix

SUBJECTIVITY IN VALIDATION RESULTS LIMITS SIMULATION APPLICABILITY

The amount of subjectivity and associated uncertainty contained in present day validation assessments significantly limit their reliability and, therefore, the confidence in the results from any simulation validated by them.

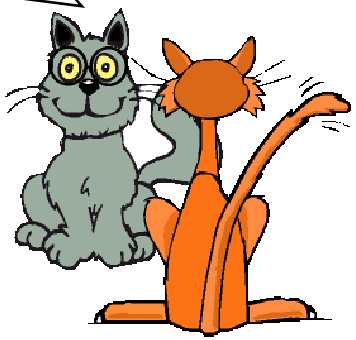

So, how's my bird simulation coming?
Well, the simulation successfully generates bird responses to a cat …
but the subjectivity in my validation assessment prevents my knowing if that's close to what a real bird will do. What's worse, I can't even say how much more I'll need to do to get the predictive accuracy that you want.
What?! How many cat treats are you charging me for this?
prediction accuracy
0
development effort

The reliability of validation assessments becomes increasingly important as a user's dependence upon a simulation's predictive value increases.

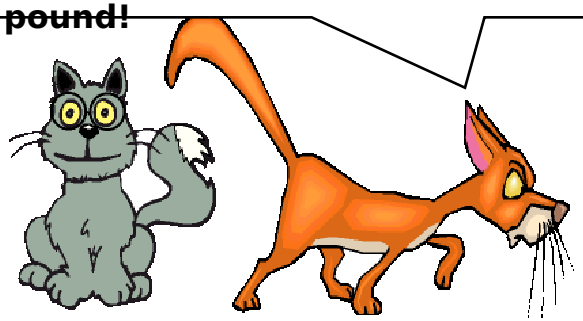
Zetetix

LACK OF RELIABLE VALIDATION ASSESSMENTS SEVERELY HAMPERS ESTIMATING COSTS

1. Without knowing how much more work is needed, I can't make any defensible cost estimates.

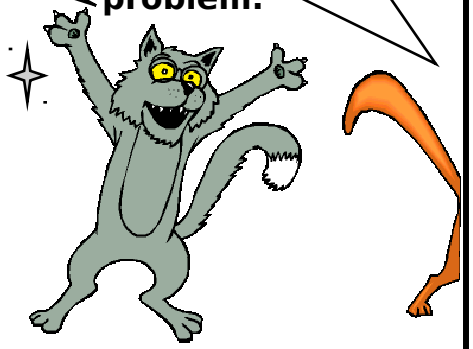


2. You'd better fix these problems or your next job will be mousing at the dog pound!



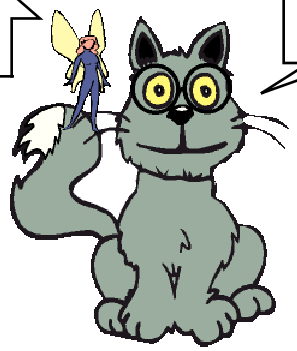
3. Wait! But, how?

You're the smart cat. That's your problem.



4. You need to know about objective validation.

What's that?



Reliable validation assessments create the crucial link between acceptability criteria and the development effort and costs because it clearly indicates the distance between what has been developed and what the user needs.

Zetetix

OBJECTIVE VALIDATION CAN BROADLY IMPROVE VALIDATION RELIABILITY & UTILITY

Objective validation reduces the subjectivity in the validation information and, ultimately, in the final validation assessments by

- **Employing well defined and repeatable processes for deriving validation information and results;**
- **Making reliable validity assessments available earlier in the development process through objective design validation,**
- **Measuring and describing validation information in directly comparable and, where possible, quantitative, terms;**
- **Characterizing the error sources in validation information, estimating the magnitude of those error sources, and assessing the importance of those error sources to validation assessment correctness;**
- **Regarding any necessary subjectivity as sources of uncertainty and quantitatively characterizing the variability associated that uncertainty;**
- **Guiding testing to sample simulation behavior space so as to efficiently gain the data most relevant to validation, and**
- **Building fundamental knowledge about simulations, their uses and their validity for those purposes.**

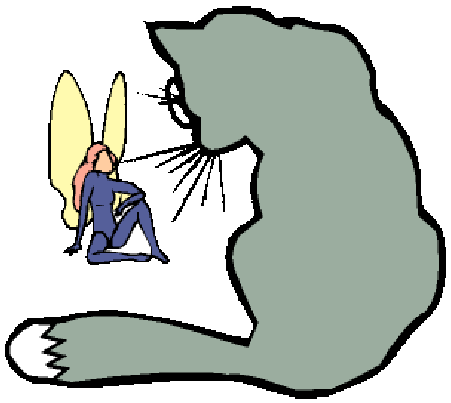
Despite impressive advances in other areas, the process of developing simulations remains largely subjective from the definition of their requirements until the final evaluation of their results to determine their fitness for purpose.

Zetetix

FIVE PROJECTS INITIATE THE DRIVE TOWARD OBJECTIVE VALIDATION

The proposed research initiative into objective validation begins with five projects aimed at reducing the largest sources of subjectivity in today's validation assessments:

- **Improving representational acceptability criteria derivation**
- **Objectively characterizing SME-supplied referents**
- **Consistently describing simulation requirements and capabilities**
- **Discovering simulation capabilities from existing software**
- **Tailoring results collection efforts to improve validation reliability**



Each of these projects targets one or more of the previously noted va

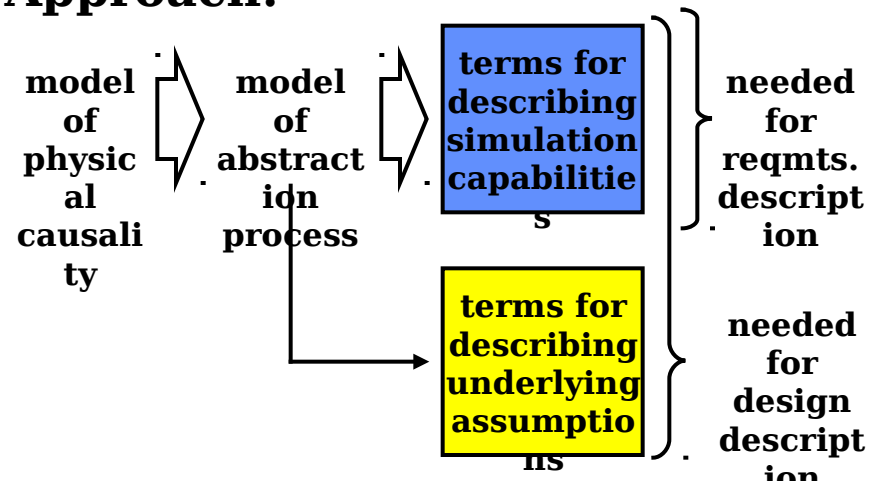
Zetetix

CONSISTENTLY DESCRIBING SIMULATION REQUIREMENTS AND CAPABILITIES

Problem: No common language exists to describe either required or supplied representational capabilities. This prevents unambiguous comparison of requirements against capabilities & complicates definition of simulation referents.

Objective: Develop a consistent & formal language for describing simulation representational capabilities.

Approach:


model of physical causality
model of abstraction process
terms for describing simulation capabilities
terms for describing underlying assumptions
needed for reqmts. description
needed for design description

Approach:

- Describe simuland behaviors & their causal relationships with causal & state graphs
- Identify the various operations for transforming simulation causal & state graphs into executable abstractions
- Characterize the mathematical properties & underlying assumptions of each abstraction operation
- Assemble these descriptions into consistent terms for describing simulation capabilities & design assumptions

Research Issues:

- Developing concise representations of non-metric property spaces
- Identifying a provably complete set of abstraction techniques
- Constructing a single mathematically consistent formalism for describing disparate abstraction operations

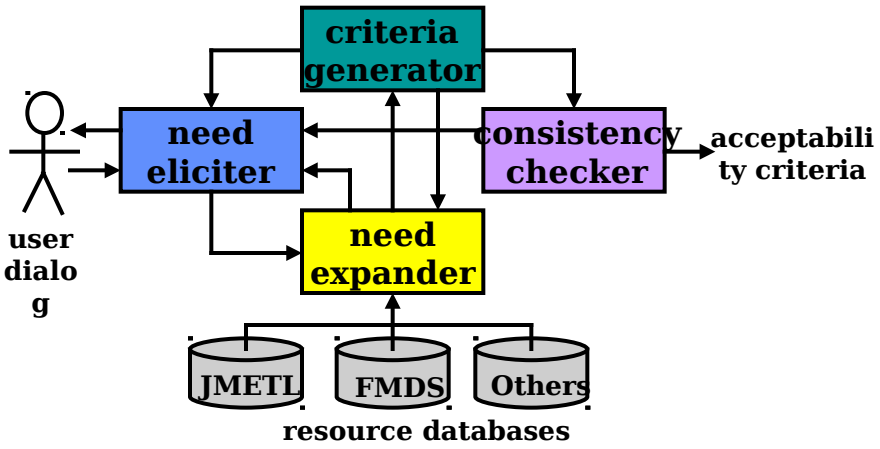
Zetetix

IMPROVING REPRESENTATIONAL ACCEPTABILITY CRITERIA DERIVATION

Problem: Users typically define very loose requirements for simulation. Every succeeding validation step depends upon the strength of these requirements.

Objective: Develop the techniques to rigorously describe representational acceptability criteria for a simulation from informal user need statements.

Approach:



Approach:

- Interactively expand user need abstractions into detailed acceptability criteria expressed as required simulation capabilities
- Derive abstraction fan-out with information from task & functional description data bases
- Apply knowledge-based system & knowledge elicitation theory to develop rules for choosing consistent criteria
- Build taxonomy of requirements that links abstract needs to details in the data bases

Research Issues:

- Coupling efficiently & effectively to information in existing data bases
- Constructing a sufficiently rich requirements abstraction taxonomy to be useful
- Integrating diverse theories on knowledge base construction & consistency into adequate guidance

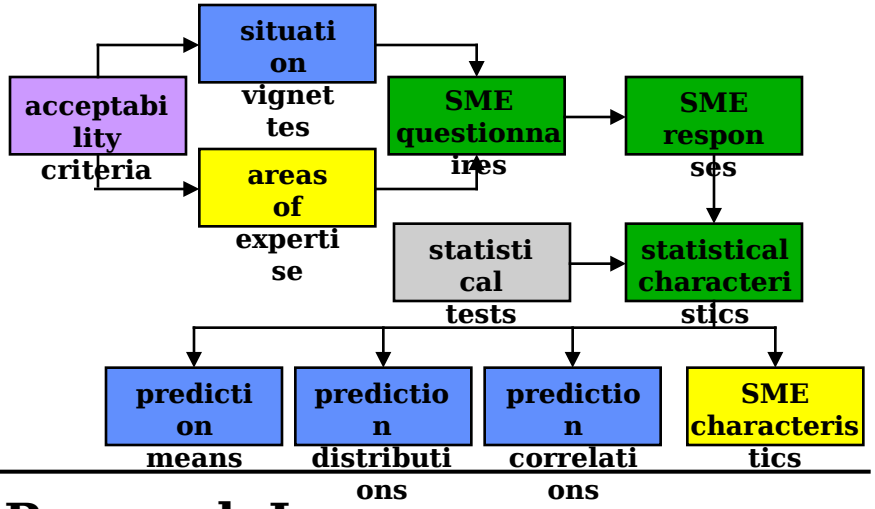
Zetetix

OBJECTIVELY CHARACTERIZING SME SUPPLIED REFERENTS

Problem: Subject matter experts (SMEs) will provide the referent knowledge for complex simulations for sometime to come. Unfortunately, SME knowledge has proven extremely hard to characterize objectively.

Objective: Develop quantitative techniques to derive simulation referents from SME knowledge.

Approach:



Approach:

- Construct situation vignettes that provide overlapping coverage of required representational space
- Formulate & administer questionnaires that sample SME predictions & expertise for each vignette
- Compute prediction means & distribution characteristics from the questionnaire responses
- Apply statistical techniques to quantify bias & identify correlations
- Describe referent by prediction means, distribution characteristics & expected

Research Issues:

- Selecting the appropriate survey techniques to minimize error sources
- Identifying & characterizing all important error sources
- Dealing with qualitative predictions
- Identifying causal relationships underlying statistically significant correlations

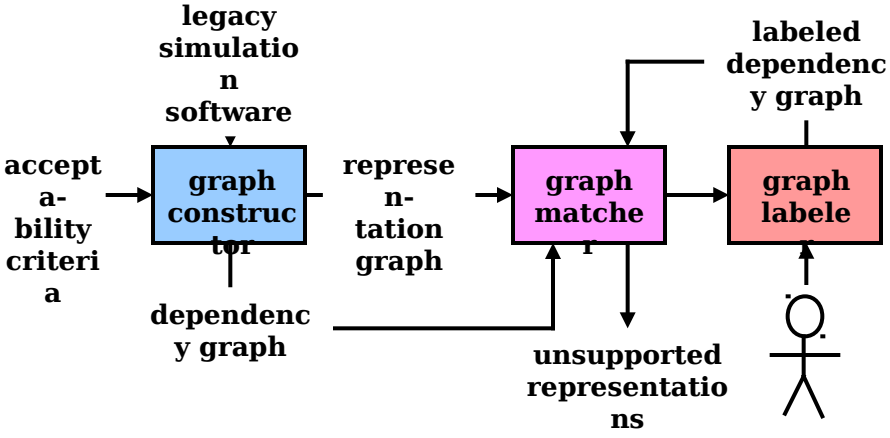
Zetetix

DISCOVERING SIMULATION CAPABILITIES FROM EXISTING SOFTWARE

Problem: The lack of design documentation describing legacy simulation capabilities makes evaluating the potential of their reuse difficult to impossible.

Objective: Develop techniques to characterize existing simulation representational capabilities from the simulation's software.

Approach:



Approach:

- Construct causal graph of required representational capabilities from acceptability criteria
- Construct dependency graph from legacy software
- Match the causal graph to the dependency graph
- Interactively label dependency graph with representation information
- Complete construction of labeled dependency graph
- Determine if causal graph maps completely into the dependency graph

Research Issues:

- Constructing meaningfully complete dependency graph from software input
- Interactive labeling the dependency graph with acceptable human workload
- Mapping causal graphs into dependency graphs with acceptable computational workload
- Deciphering multiple partial mappings
- Handling parameter input effects upon representational capabilities

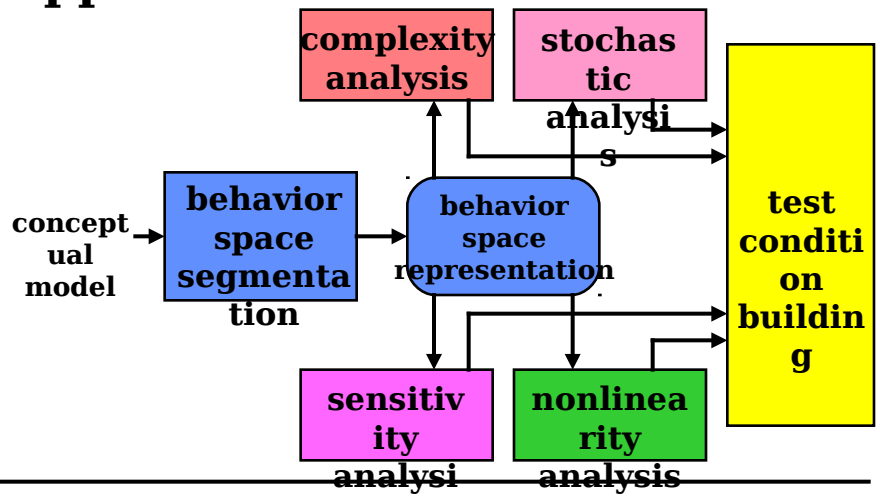
Zetetix

TAILORING RESULTS COLLECTION EFFORTS TO IMPROVE VALIDATION RELIABILITY

Problem: The complicated simulations of today generate extremely complex behavior spaces, spaces that existing testing techniques alone cannot reliably characterize through test sampling.

Objective: Develop techniques that guide testing sampling of a complex simulation's behavior space from the output of previous validation steps.

Approach:



Approach:

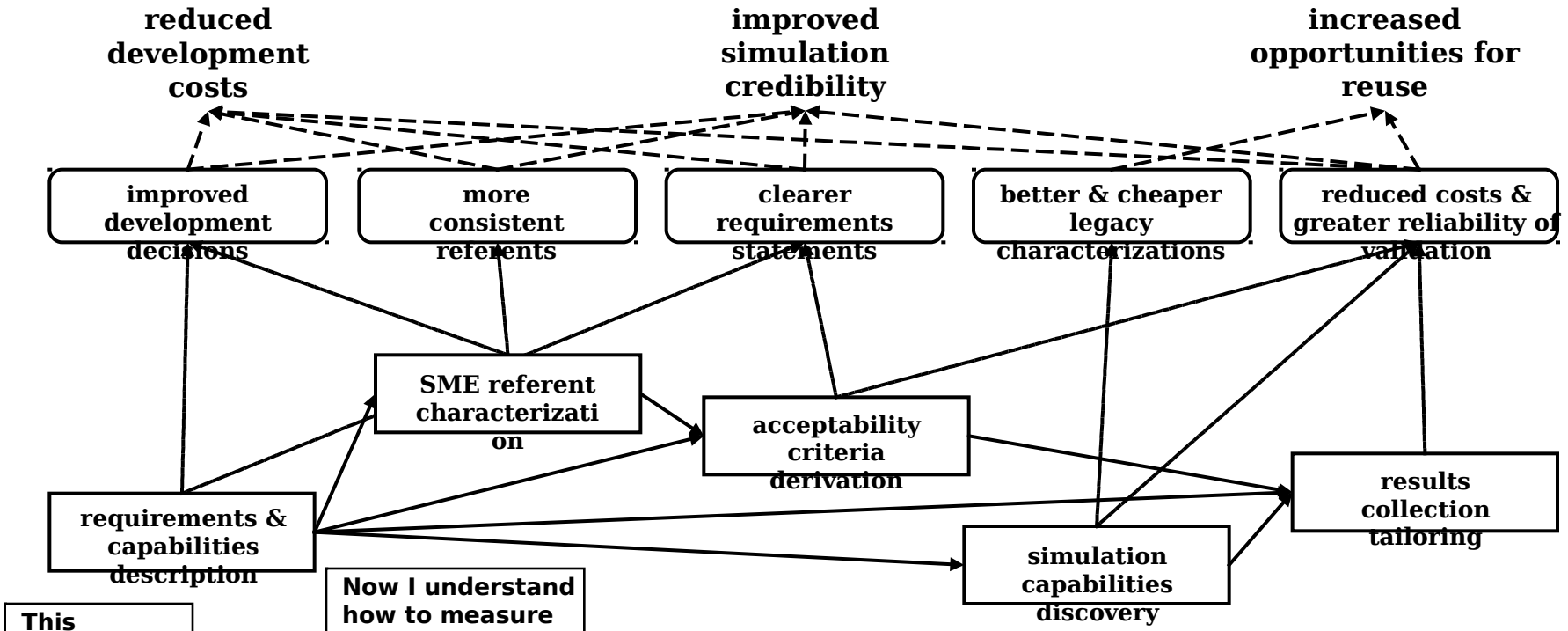
- Analyze conceptual model to identify possible nonlinearity, stochasticity & chaos in behavior space
- Perform sensitivity analysis within dependency domains to identify areas possibly important to purpose
- Estimate required sample sizes to needed to sufficiently characterize behavior
- Recommend test conditions & procedures that adequately sample likely problem spaces

Research Issues:

- Analyzing conceptual model representation for simulation behavior properties
- Performing sensitivity analysis with acceptable computational loading
- Identifying meaningful & accessible test conditions from analysis results

Zetetix

ROADMAP GUIDES RESEARCH TO REALIZE IMPROVEMENTS TO VALIDATION PRACTICE

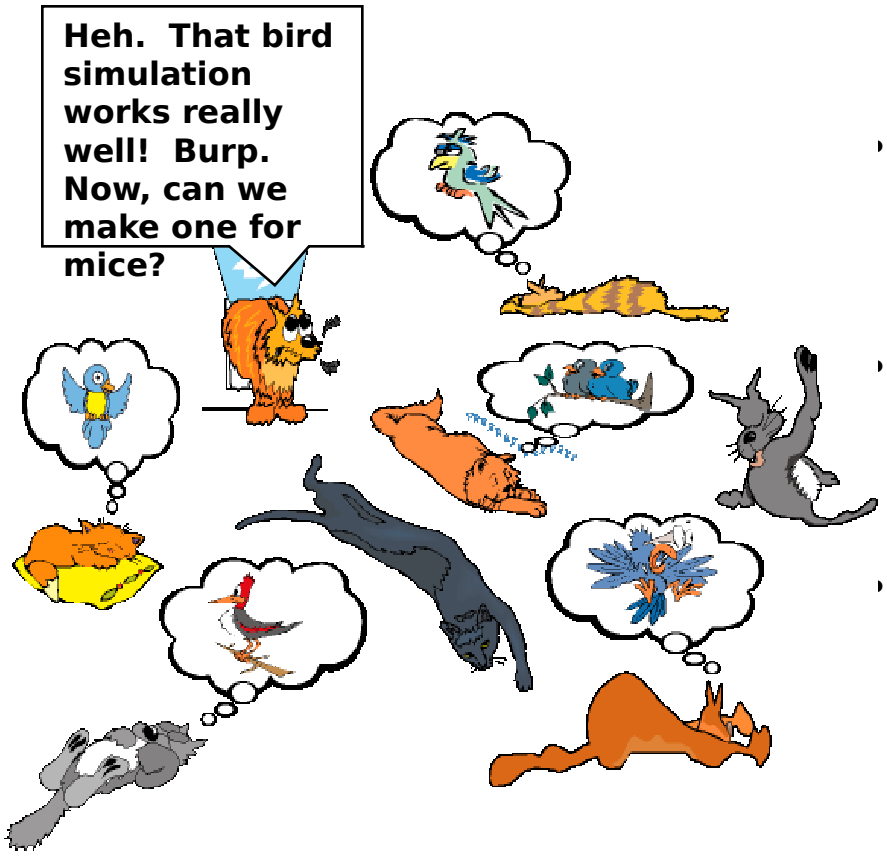

reduced development costs
improved simulation credibility
increased opportunities for reuse
improved development decisions
more consistent referents
clearer requirements statements
better & cheaper legacy characterizations
reduced costs & greater reliability of validation
SME referent characterization
acceptability criteria derivation
requirements & capabilities description
results collection tailoring
simulation capabilities discovery
This research program will take you to where you need to be.
Now I understand how to measure and use validity in that bird simulation.

The individual research projects build upon one another to progressively improve validation practice that will, ultimately, realize the simulation-wide goals of reduced costs, improved credibility and increased reuse.

Zetetix

SUMMARY & CONCLUSIONS

Some time later …



Morale: The proof of simulation validity is the eating that it makes possible.

- Validation determines how much simulation is needed to achieve a purpose.
- Validation of today's simulations involves considerable subjectivity that weakens the reliability of their assessments.
- The proposed research program on objective validation will drive the subjectivity out of existing validation practice.
- Objective validation will make validation assessments more reliable, consistent, repeatable, trustworthy and useful thus making highly predictive simulations more possible and improving people's choices when using existing simulations.

Zetetix